

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 16/573,381  
Source: IFWO  
Date Processed by STIC: 3/6/07

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 03/06/2007

PATENT APPLICATION: US/10/573,381

TIME: 11:17:28

Input Set : N:\efs\03\_06\_07\10573381\_efs\546572Seq.txt

Output Set: N:\CRF4\03062007\J573381.raw

```

3 <110> APPLICANT: TAKARA BIO INC.
5 <120> TITLE OF INVENTION: Polypeptide having RNaseIII activity
7 <130> FILE REFERENCE: 664746
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/573,381
C--> 9 <141> CURRENT FILING DATE: 2006-03-24
9 <150> PRIOR APPLICATION NUMBER: JP 2003-342260
10 <151> PRIOR FILING DATE: 2003-09-30
12 <150> PRIOR APPLICATION NUMBER: JP 2003-409638
13 <151> PRIOR FILING DATE: 2003-12-08
15 <160> NUMBER OF SEQ ID NOS: 17
17 <170> SOFTWARE: PatentIn version 3.1
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 678
21 <212> TYPE: DNA
22 <213> ORGANISM: Shewanella sp.AC10
24 <400> SEQUENCE: 1
25 atggaaccca ttaaaaattt gccgcgtttg tgccgtactt tagggttatga gttcaataat      60
27 attgaattac ttattcaggc cttaacacat cgtagcgcag caaataaaca taatgagcgt      120
29 ttagagtttt taggtgattc gattttatcg atagccattt cagatgcctt atatcatcag      180
31 tttccaaagg cgactgaagg tgatttaagc cgaatgcgcg ccactttagt caaagggtgac      240
33 acgctgacaa tcatagctaa agagttcaag ctaggtgatt atttgtattt aggtccttgg      300
35 gaactcaaaa gtggtggctt tagacgcgaa tctatttttag ctgatgctgt agaggctatt      360
37 attggtgctg tctatcttga tgctgatatt gaagtgtgcc gcaagctatt attatcatgg      420
39 tatcaagagc gtttagctga gattaaaccg ggtattaatc aaaaagatcc gaagacaata      480
41 ttgcaagaat acctgcaagg ttttaaaaag ccattgcctg attaccaagt tgttgacagta      540
43 gaagggtgaag cccatgatca aaccttcacc gtagaatgta aaattagtag attagataaa      600
45 gttgtcaccg gtgtggcaag ttcaagaaga aaagctgaac agcttgccgc tgctcaggta      660
47 ttggagctac tgaataaa
50 <210> SEQ ID NO: 2
51 <211> LENGTH: 39
52 <212> TYPE: DNA
53 <213> ORGANISM: Artificial
55 <220> FEATURE:
56 <223> OTHER INFORMATION: Synthetic primer 1 to amplify a gene encoding Shewanella
sp.AC10 RNaseIII
58 <400> SEQUENCE: 2
59 cagattccac gaattcgatg gaaccatta aaaatttgc      39
62 <210> SEQ ID NO: 3
63 <211> LENGTH: 37
64 <212> TYPE: DNA
65 <213> ORGANISM: Artificial
67 <220> FEATURE:
68 <223> OTHER INFORMATION: Synthetic primer 2 to amplify a gene encoding Shewanella
sp.AC10 RNaseIII

```

70 <400> SEQUENCE: 3

## RAW SEQUENCE LISTING

DATE: 03/06/2007

PATENT APPLICATION: US/10/573,381

TIME: 11:17:28

Input Set : N:\efs\03\_06\_07\10573381\_efs\546572Seq.txt

Output Set: N:\CRF4\03062007\J573381.raw

37

71 ggagaggtct ggatccttat ttattcagta gctcctt

74 &lt;210&gt; SEQ ID NO: 4

75 &lt;211&gt; LENGTH: 226

76 &lt;212&gt; TYPE: PRT

77 &lt;213&gt; ORGANISM: Shewanella sp.Ac10

79 &lt;400&gt; SEQUENCE: 4

```

81 Met Glu Pro Ile Lys Asn Leu Pro Arg Leu Cys Arg Thr Leu Gly Tyr
82 1          5          10          15
85 Glu Phe Asn Asn Ile Glu Leu Leu Ile Gln Ala Leu Thr His Arg Ser
86          20          25          30
89 Ala Ala Asn Lys His Asn Glu Arg Leu Glu Phe Leu Gly Asp Ser Ile
90          35          40          45
93 Leu Ser Ile Ala Ile Ser Asp Ala Leu Tyr His Gln Phe Pro Lys Ala
94          50          55          60
97 Thr Glu Gly Asp Leu Ser Arg Met Arg Ala Thr Leu Val Lys Gly Asp
98 65          70          75          80
101 Thr Leu Thr Ile Ile Ala Lys Glu Phe Lys Leu Gly Asp Tyr Leu Tyr
102          85          90          95
105 Leu Gly Pro Gly Glu Leu Lys Ser Gly Gly Phe Arg Arg Glu Ser Ile
106          100          105          110
109 Leu Ala Asp Ala Val Glu Ala Ile Ile Gly Ala Val Tyr Leu Asp Ala
110          115          120          125
113 Asp Ile Glu Val Cys Arg Lys Leu Leu Leu Ser Trp Tyr Gln Glu Arg
114          130          135          140
117 Leu Ala Glu Ile Lys Pro Gly Ile Asn Gln Lys Asp Pro Lys Thr Ile
118 145          150          155          160
121 Leu Gln Glu Tyr Leu Gln Gly Phe Lys Lys Pro Leu Pro Asp Tyr Gln
122          165          170          175
125 Val Val Ala Val Glu Gly Glu Ala His Asp Gln Thr Phe Thr Val Glu
126          180          185          190
129 Cys Lys Ile Ser Glu Leu Asp Lys Val Val Thr Gly Val Ala Ser Ser
130          195          200          205
133 Arg Arg Lys Ala Glu Gln Leu Ala Ala Ala Gln Val Leu Glu Leu Leu
134          210          215          220
137 Asn Lys
138 225

```

141 &lt;210&gt; SEQ ID NO: 5

142 &lt;211&gt; LENGTH: 243

143 &lt;212&gt; TYPE: PRT

144 &lt;213&gt; ORGANISM: Artificial

146 &lt;220&gt; FEATURE:

147 &lt;223&gt; OTHER INFORMATION: An expression peptide sequence of Shewanella sp.AC10

RNaseIII

149 &lt;400&gt; SEQUENCE: 5

```

151 Met Asn His Lys Val His His His His His Ile Glu Gly Arg Asn
152 1          5          10          15
155 Ser Met Glu Pro Ile Lys Asn Leu Pro Arg Leu Cys Arg Thr Leu Gly
156          20          25          30
159 Tyr Glu Phe Asn Asn Ile Glu Leu Leu Ile Gln Ala Leu Thr His Arg
160          35          40          45

```

## RAW SEQUENCE LISTING

DATE: 03/06/2007

PATENT APPLICATION: US/10/573,381

TIME: 11:17:28

Input Set : N:\efs\03\_06\_07\10573381\_efs\546572Seq.txt

Output Set: N:\CRF4\03062007\J573381.raw

```

163 Ser Ala Ala Asn Lys His Asn Glu Arg Leu Glu Phe Leu Gly Asp Ser
164      50                      55                      60
167 Ile Leu Ser Ile Ala Ile Ser Asp Ala Leu Tyr His Gln Phe Pro Lys
168 65                      70                      75                      80
171 Ala Thr Glu Gly Asp Leu Ser Arg Met Arg Ala Thr Leu Val Lys Gly
172                      85                      90                      95
175 Asp Thr Leu Thr Ile Ile Ala Lys Glu Phe Lys Leu Gly Asp Tyr Leu
176                      100                     105                     110
179 Tyr Leu Gly Pro Gly Glu Leu Lys Ser Gly Gly Phe Arg Arg Glu Ser
180                      115                     120                     125
183 Ile Leu Ala Asp Ala Val Glu Ala Ile Ile Gly Ala Val Tyr Leu Asp
184                      130                     135                     140
187 Ala Asp Ile Glu Val Cys Arg Lys Leu Leu Leu Ser Trp Tyr Gln Glu
188 145                      150                      155                      160
191 Arg Leu Ala Glu Ile Lys Pro Gly Ile Asn Gln Lys Asp Pro Lys Thr
192                      165                      170                      175
195 Ile Leu Gln Glu Tyr Leu Gln Gly Phe Lys Lys Pro Leu Pro Asp Tyr
196                      180                      185                      190
199 Gln Val Val Ala Val Glu Gly Glu Ala His Asp Gln Thr Phe Thr Val
200                      195                      200                      205
203 Glu Cys Lys Ile Ser Glu Leu Asp Lys Val Val Thr Gly Val Ala Ser
204                      210                      215                      220
207 Ser Arg Arg Lys Ala Glu Gln Leu Ala Ala Ala Gln Val Leu Glu Leu
208 225                      230                      235                      240

```

211 Leu Asn Lys

215 &lt;210&gt; SEQ ID NO: 6

216 &lt;211&gt; LENGTH: 720

217 &lt;212&gt; TYPE: DNA

218 &lt;213&gt; ORGANISM: Artificial

220 &lt;220&gt; FEATURE:

221 &lt;223&gt; OTHER INFORMATION: red-shifted green fluorescence protein

223 &lt;400&gt; SEQUENCE: 6

```

224 atggctagca aaggagaaga actcttcact ggagttgtcc caattcttgt tgaattagat      60
226 ggtgatgtta acggccacaa gttctctgtc agtggagagg gtgaagggtga tgcaacatac      120
228 ggaaaactta ccctgaagtt catctgcact actggcaaac tgccctgttcc atggccaaca      180
230 ctagtcaacta ctctgtgcta tgggtgttcaa tgcttttcaa gatacccgga tcatatgaaa      240
232 cggcatgact ttttcaagag tgccatgccc gaaggttatg tacaggaaaag gaccatcttc      300
234 ttcaaagatg acggcaacta caagacacgt gctgaagtca agtttgaagg tgataccctt      360
236 gttaatagaa tcgagttaaa aggtattgac ttcaaggaag atggaaacat tctgggacac      420
238 aaattggaat acaactataa ctcacacaat gtatacatca tggcagacaa acaaaagaat      480
240 ggaatcaaag tgaacttcaa gaccgcgccac aacattgaag atggaagcgt tcaactagca      540
242 gaccattatc aacaaaatac tccaattggc gatggccctg tccttttacc agacaacat      600
244 tacctgtcca cacaatctgc cctttcgaaa gatcccaacg aaaagagaga ccacatggtc      660
246 cttcttgagt ttgtaacagc tgctgggatt acacatggca tggatgaact gtacaactga      720

```

249 &lt;210&gt; SEQ ID NO: 7

250 &lt;211&gt; LENGTH: 42

251 &lt;212&gt; TYPE: DNA

252 &lt;213&gt; ORGANISM: Artificial

254 &lt;220&gt; FEATURE:

## RAW SEQUENCE LISTING

DATE: 03/06/2007

PATENT APPLICATION: US/10/573,381

TIME: 11:17:28

Input Set : N:\efs\03\_06\_07\10573381\_efs\546572Seq.txt

Output Set: N:\CRF4\03062007\J573381.raw

255 <223> OTHER INFORMATION: Synthetic primer dsr-1 to amplify a gene encoding red-shifted green fluorescence protein

257 <400> SEQUENCE: 7

258 gggtaatacgc actcactata gggagaatgg ctagcaaagg ag 42

261 <210> SEQ ID NO: 8

262 <211> LENGTH: 42

263 <212> TYPE: DNA

264 <213> ORGANISM: Artificial

266 <220> FEATURE:

267 <223> OTHER INFORMATION: Synthetic primer dsr-2 to amplify a gene encoding red-shifted green fluorescence protein

269 <400> SEQUENCE: 8

270 gggtaatacgc actcactata gggagatcag ttgtacagtt ca 42

273 <210> SEQ ID NO: 9

274 <211> LENGTH: 42

275 <212> TYPE: DNA

276 <213> ORGANISM: Artificial

278 <220> FEATURE:

279 <223> OTHER INFORMATION: Synthetic primer dsl-1 to amplify a gene encoding luciferase

281 <400> SEQUENCE: 9

282 gggtaatacgc actcactata gggagaatgg aagacgccaa aa 42

285 <210> SEQ ID NO: 10

286 <211> LENGTH: 42

287 <212> TYPE: DNA

288 <213> ORGANISM: Artificial

290 <220> FEATURE:

291 <223> OTHER INFORMATION: Synthetic primer dsl-2 to amplify a gene encoding luciferase

293 <400> SEQUENCE: 10

294 gggtaatacgc actcactata gggagagaac gtgtacatcg ac 42

297 <210> SEQ ID NO: 11

298 <211> LENGTH: 42

299 <212> TYPE: DNA

300 <213> ORGANISM: Artificial

302 <220> FEATURE:

303 <223> OTHER INFORMATION: Synthetic primer dsl-3 to amplify a gene encoding luciferase

305 <400> SEQUENCE: 11

306 gggtaatacgc actcactata gggagaggca gatggaacct ct 42

309 <210> SEQ ID NO: 12

310 <211> LENGTH: 66

311 <212> TYPE: PRT

312 <213> ORGANISM: Thermotoga maritima

314 <400> SEQUENCE: 12

316 Met Arg Gly Lys Val Lys Trp Phe Asp Ser Lys Lys Gly Tyr Gly Phe

317 1 5 10 15

320 Ile Thr Lys Asp Glu Gly Gly Asp Val Phe Val His Trp Ser Ala Ile

321 20 25 30

324 Glu Met Glu Gly Phe Lys Thr Leu Lys Glu Gly Gln Val Val Glu Phe

325 35 40 45

328 Glu Ile Gln Glu Gly Lys Lys Gly Pro Gln Ala Ala His Val Lys Val

329 50 55 60

332 Val Glu

## RAW SEQUENCE LISTING

DATE: 03/06/2007

PATENT APPLICATION: US/10/573,381

TIME: 11:17:28

Input Set : N:\efs\03\_06\_07\10573381\_efs\546572Seq.txt

Output Set: N:\CRF4\03062007\J573381.raw

```

333 65
336 <210> SEQ ID NO: 13
337 <211> LENGTH: 198
338 <212> TYPE: DNA
339 <213> ORGANISM: Thermotoga maritima
341 <400> SEQUENCE: 13
342 atgagaggaa aggttaagtg gttcgattcc aagaagggtc acggattcat cacaaaggac      60
344 gaaggaggag acgtgttcgt aacttggtca gccatcgaaa tggaagggtt caaaactctg      120
346 aaggaaggcc aggtcgtcga gttcgagatt caggaaggca agaaagggtc acaggcagcg      180
348 cacgtgaaaag tagttgag                                     198
350 <210> SEQ ID NO: 14
351 <211> LENGTH: 20
352 <212> TYPE: DNA
353 <213> ORGANISM: Artificial
355 <220> FEATURE:
356 <223> OTHER INFORMATION: Synthetic primer rsGFP-F to amplify a gene encoding rsGFP
358 <400> SEQUENCE: 14
359 gccacaacat tgaagatgga                                     20
362 <210> SEQ ID NO: 15
363 <211> LENGTH: 20
364 <212> TYPE: DNA
365 <213> ORGANISM: Artificial
367 <220> FEATURE:
368 <223> OTHER INFORMATION: Synthetic primer rsGFP-R to amplify a gene encoding rsGFP
370 <400> SEQUENCE: 15
371 gaaagggcag attgtgtgga                                     20
373 <210> SEQ ID NO: 16
374 <211> LENGTH: 20
375 <212> TYPE: DNA
376 <213> ORGANISM: Artificial
378 <220> FEATURE:
379 <223> OTHER INFORMATION: Synthetic primer Neo-F to amplify a gene encoding Neo
381 <400> SEQUENCE: 16
382 atagcgttgg ctacccgtga                                     20
385 <210> SEQ ID NO: 17
386 <211> LENGTH: 20
387 <212> TYPE: DNA
388 <213> ORGANISM: Artificial
390 <220> FEATURE:
391 <223> OTHER INFORMATION: Synthetic primer Neo-R to amplify a gene encoding Neo
393 <400> SEQUENCE: 17
394 gaaggcgata gaaggcgatg                                     20

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 03/06/2007  
PATENT APPLICATION:    US/10/573,381      TIME: 11:17:29

Input Set : N:\efs\03\_06\_07\10573381\_efs\546572Seq.txt  
Output Set: N:\CRF4\03062007\J573381.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:2; Line(s) 56  
Seq#:3; Line(s) 68  
Seq#:7; Line(s) 255  
Seq#:8; Line(s) 267

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,  
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:2,3,5,6,7,8,9,10,11,14,15,16,17



VERIFICATION SUMMARY

DATE: 03/06/2007

PATENT APPLICATION: US/10/573,381

TIME: 11:17:29

Input Set : N:\efs\03\_06\_07\10573381\_efs\546572Seq.txt

Output Set: N:\CRF4\03062007\J573381.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No

L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date